

## Chapter 12

# I Mourned the University for a Long Time: Michael Sperberg-McQueen and Julianne Nyhan

**Abstract** This interview took place on 9 July 2014 at dh2014, the Digital Humanities Conference that was held in Lausanne, Switzerland that year. In it Sperberg-McQueen recalls having had some exposure to programming in 1967, as a 13 year-old. His next notable encounter with computing was as a graduate student when he set about using computers to make a bibliography of secondary literature on the Elder Edda. His earliest encounters with Humanities Computing were via books, and he mentions the proceedings of the ‘Concordances and the Dictionary of Old English’ conference and a book by Susan Hockey (see below) as especially influential on him. In 1985 a position in the Princeton University Computer Center that required an advanced degree in Humanities and knowledge of computing became available; he took on the post while finishing his PhD dissertation and continuing to apply for tenure-track positions. Around this time he also began attending the ‘International Conference on Computers and the Humanities’ series and in this interview he describes some of the encounters that took place at those conferences and contributed to the formation of projects like TEI. As well as reflecting on his role in TEI he also compares and contrasts this experience with his work in W3C. On the whole, a somewhat ambivalent attitude towards his career emerges from the interview: he evokes Dorothy Sayers to communicate how the application of computers to the Humanities ‘overmastered’ him. Yet, he poignantly recalls how his first love was German Medieval languages and literature and the profound sense of loss he felt at not securing an academic post related to this.

## Biography

**Michael Sperberg McQueen** was born in 1954 in Borger, northern Texas. At present he is Principal of Black Mesa Technologies, a limited liability company that specialises in XML and other descriptive markup technologies. His PhD (1985) from Stanford University is in Comparative Literature. From 1988 to 2000 he was the Editor in Chief of the TEI; from 1996 to 1998 he served as co-Editor of the Extensible Markup Language (XML) 1.0 specification. He was also a member of the technical staff of the World Wide Web Consortium (W3C) from 1998 to 2009 and performed various functions in this role including staff contact of the W3C XML Schema Working Group and co-editor of the XSD 1.1 specification. He has

been a visiting researcher at the University of Bergen and, more recently, a visiting Professor in the Program in DH, Dept. of Linguistics and Literary Studies, Technical University of Darmstadt (*Institut für Sprach- und Literaturwissenschaft, Technische Universität Darmstadt*). As well as his internationally acknowledged work on XML (which has become the lingua franca of data structure and exchange in very many domains) his scholarship on knowledge formalisation is of seminal importance to DH where the TEI has become the de facto standard for making Humanities texts machine readable.

## Interview

**JN** My first question is about your earliest memory, in any context at all, of encountering computing or computing technology?

**MSMQ** My first direct encounter with computing technology was in the Summer of 1967. I think I must have been 13, and some programme in the public schools offered, I think, a programming course. It was offered through some programme that I was involved with and a friend of mine and I said “ok, we’ll go to this programming course”. I went for a couple of weeks but then my friend didn’t want to go anymore and I couldn’t get a ride so I stopped. But I had 2 weeks of exposure to FORTRAN and they started by giving a test to distinguish people who had already learned a bit of programming from people who didn’t.

They obviously hadn’t instrumented it very well because they asked questions about what parenthesised expressions would mean, and whether in  $a+b*x$  the multiplication or the addition would bind more tightly. I’d had no computing experience at all but just the sense of the expression was obvious and they said “oh, you must have had programming”. So they put me in the advanced class and I couldn’t figure out anything because I didn’t know anything about computers or FORTRAN.

After that, as they say “*lange Zeit gar nichts*” [nothing for a long time]. The next contact would have been as a graduate student. Well, sorry, occasional things, computing, punch cards. One had periodically at that point in the 1960s, 1970s and 1980s, during my school and university time, encounters with organisations that used punch cards and so forth for organisation.

**JN** Why did you want to take the course in programming? That first course?

**MSMQ** It sounded intellectually challenging, I think. I don’t remember more than that.

**JN** That’s very impressive for a 13 year old!

**MSMQ** I think I was in seventh or eighth grade. The next time I remember thinking at all seriously about computing was, I believe, as a graduate student. I was a Medievalist and I ran across a collection of essays, actually it was the Proceedings of a small conference held at the University of Toronto called 'Concordances and the Dictionary of Old English'. It was a planning conference that the *Dictionary of Old English* (DOE) people had organised to talk about how computers might help them write a new DOE (see Cameron et al. 1970). As a Medievalist I had spent a lot of time, as everyone I knew in Medieval Studies or Classics did, transcribing glossary entries on to index cards and transcribing locations of occurrences of words on to index cards so I could sort them and re-sort them and analyse them and think "what's the meaning of this word as opposed to that word? How many different words are used for 'King' in Beowulf? What are the nuances of the different words? What's their etymology? And so forth".

So you spent a lot of time leafing through Klaeber's glossary (1936) and the idea that you could generate a concordance automatically seemed like magic. I remember talking to other people about it and mentioning it to my advisers. One adviser said, "I wouldn't get involved with that if I were you" and I said, "why not? It seems like the obvious thing to do, it seems like the way to build better tools for Medieval Studies". He said, "yeah, but everybody who gets involved in computers are pretty soon spending all their time doing computer stuff and not Philology". I always thought that in later years he must have told his students the same thing and pointed to me as an awful example: "he's never gotten a job in Philology", as indeed was the case. The other adviser, on the other hand, handed me a shoe box and said "this is the bibliography of the Elder Edda [Old Norse poems that are primary sources for Scandinavian mythology and heroic legend] since 1953. The goal of this project is to computerise it and your job is to figure out what that would mean and then do it". I learned a lot, I made a lot of stupid, ignorant mistakes and I learned a lot from my stupid, ignorant mistakes.

**JN** Did you have some access to formal training by that point?

**MSMQ** When he handed me that shoe box I went down to the computer center and signed up for their 'Introduction to the Computer Center' course and all the other courses that seemed relevant. Of course, as you go through one course you learn about other things that are relevant and so I had the kind of short course training (3 or 4 hours), that was offered by computer centers at that point. It may still be offered by some computer centers somewhere, although probably not so much anymore. But I have never had any formal academic training in computing.

**JN** And have you been for the most part self-taught? What sort of strategies did you use?

**MSMQ** As a beginning user, first at Stanford, where I was at graduate school, and then at Johns Hopkins University, where my wife got a job and I had access to computing, I was still finishing my dissertation. I was using mainframes and because we

didn't own a terminal or a modem, using a mainframe meant that you had to go to a terminal room on campus. And if you go to a terminal room often enough you see who's there all the time and you can get some notion, just by glancing at their screens as you walk past them to a free space, what kind of thing they're doing. And you overhear people talking and so forth and eventually you get some notion in a completely informal way of people who are sharp and who may help. And if you walk past somebody and they do something clever with the system editor, you can say, "wait, how did you do that?" They're often happy to show you. So, in fact, a lot of the practical interaction with computing seemed to me to be conveyed through a kind of oral tradition. You could learn parts of it by reading documentation but I spent a lot of time studying documentation and I found a lot of it completely impenetrable because it was not in my vocabulary. So I learned a lot by looking at other people and from that sort of informal helping. I often have wondered, how do people learn that kind of thing now, when they don't have to go to a terminal room? In some universities, I guess, PC pools still exist and presumably still have similar social effects but I don't know.

**JN** I think it's an interesting question. I wonder, especially as DH gets more established and formalised, about the types of differences in modes of learning that will follow, and the implications of this.

**MSMQ** Of course, at some point I did set out to teach myself computing more seriously, in particular, in 1985 when I got my first job at the Princeton University Computing Service. I said "oh my God, I'm in a computer center, I've got to learn about computers, I'm responsible for advising Humanists on the use of computers, I have to understand this." So I spent a lot of time going to the library and reading about databases and compilers and so forth. And compilers always seemed interesting, partly because they were magic and partly because they involved something called parsing and that sounded like language processing and that sounded interesting. So I have, I believe, many of the odd unevennesses in my knowledge that you find with some autodidacts because they go very deep in some areas and they are completely ignorant about some other things.

**JN** So, you seem to indicate that you didn't get a job in Philology because you had pursued computing to the extent that you did. Is that interpretation correct?

**MSMQ** The causality could be, is probably, a far step. But it is true that, as I normally put it to myself, I never got a job. Certainly as I conceived of the world as a graduate student, no job I've ever had has counted as a job. I never had a teaching job.

**JN** Because you were forced on this professional route?

**MSMQ** Yes, yes, the only reason to study Old Norse as far as I could tell was to become a professor of Old Norse.

**JN** Will you talk a little bit about the process of looking for ‘the job’, so to speak, and how it was that you ended up in the Computer Center and your emotions and thoughts about that?

**MSMQ** Sure. I was finishing my dissertation and starting to look for jobs. At the time I was finishing my dissertation or doing my dissertation universities in the US and Canada were producing probably, judging from *Dissertation Abstracts*, I think there were on average 10 Medieval Germanists a year, give or take, including Old Norse. And there were maybe two or three tenure track jobs that mentioned Medieval German as a potential area of specialisation. So I had friends in graduate school who applied to every position in English or the language that could possibly fit for them and they were sending out 200 applications. None of the Germanists could find 200 institutions to write to, so the chances were very great that the large majority of people getting PhDs in the kind of field that I was in were not going to get academic jobs.

Of course, I always expected to be the exception. I was finishing working on my dissertation and a friend of ours who had done her degree at Johns Hopkins, where my wife was teaching, phoned us one Sunday. She said “have you read the *New York Times* today? Have you looked at the Education Supplement?” And we said “no, we haven’t got it today” and she said, “stop what you’re doing, go out and get the *Times*, get the Education Supplement and turn to page 13. There is a job there with your name on it.” I said “ok.” We went out and got a copy of the *New York Times* and the Education Supplement had an ad from Princeton University Computer Center looking for a Humanist, someone with an advanced degree in the Humanities and knowledge of computing, and it did certainly sound interesting, so I applied (in the summer of 1984, I believe).

I felt a little guilty about applying because I was very close to finishing my degree and I was applying for academic jobs. I knew that if they hired me and I started there at the computer center and then one of my academic jobs came through I would be there for 6 months and then I would leave. I felt a little bit guilty about that but I said if they ask me, maybe I’ll tell them and maybe I won’t, but if they don’t ask I’m not going to tell them. So I applied and got the job; they may have assumed more computing knowledge than I had, or I may have talked a good game, or they may have been perfectly clear that what they needed was somebody with an advanced degree in the Humanities who was willing to tolerate learning about computers and that they certainly found in me. I felt they were taking a risk but I was very grateful to them. And then, of course, the academic jobs didn’t come through. I don’t know who those people hired but that is probably just as well so I don’t resent person X or person Y!

So I found myself at the Princeton Computing Service, actually when the woman who became my boss called to offer me the job she said, “so you’re finishing your dissertation, how close are you? How long is it going to take for you to finish if you don’t come here?” I estimated a time and she said, “fine, we’ll start you after that. It’s important for me to get this decision made but it’s not important that you start

next week, you should finish before you come.” She had long experience with people being almost done. Of course, I missed my self-imposed deadline. I was probably 2 weeks away from finishing my dissertation when, in fact, the time came for me to start work and so it took 6 months because I was starting a new job and I could only work nights and I was distracted. And, of course, I thought it was 2 weeks and it was probably a little more like a month. But I finished my degree.

Then I applied for more jobs, feeling again a little guilty, but at some point your degree is old enough that the first question any research committee is going to ask (or at least this is what I thought) is “wait, he got his degree this many years ago, he’s never had an academic job, there must be something wrong with him”. You begin to look like damaged goods. So at some point I stopped applying for teaching jobs and I made my career in computing. Of course, in 1985 and 1986, the years that I was at Princeton, I attended the predecessor of the DH conference, it was the International Conference on Computers and the Humanities, it wasn’t even called ACH/ALLC yet.

**JN** Was that the first conference that you attended in this field?

**MSMQ** The first that I attended was in 1985 in Provo, Utah. I heard a talk from the President of the ACH, Nancy Ide, who I think talked about teaching programming to Humanists. Since I was trying to learn programming that sounded like an interesting topic. She said she was writing a textbook and I asked if I could see the draft of the textbook. She said “sure, on one condition”; I said, “what’s the condition?” She said “you must send me comments on every chapter that you read” and I promised to send her comments. In her book *PASCAL for the Humanities* (Ide 1987) she teaches students how to write PASCAL and the ongoing example is essentially an interactive concordance programme. And at some point she says “now, when you display the ‘hits’, the occurrences of a given word to the user, you probably want to tell them what chapter it’s in, so you need some way to tell when a new chapter starts so that you can keep your counter”. And I wrote in my comments “is there a standard way to do that because if there is you should probably mention it” (I had the importance of standardisation hammered into me at my job) “if there’s not a standard way, then isn’t this one of the things that you were asking about in the ACH General Meeting? You said “if there was anything the Association should be doing and we’re not let me know”- if there’s not a standard way for people to keep track of where chapters begin then isn’t that the kind of thing ACH should be doing?” She wrote back and said, “you know, you’re right, it is, and in fact there’s a small group of us that’s working towards some sort of text encoding format or guidelines. Would you like to be involved?” And she and I started talking about this; I never saw any of these other alleged people. I don’t think they were a fiction but I think they weren’t actually getting any forwarder, whereas Nancy Ide and I, somehow we clicked. And so, beginning in 1987, we were working on the TEI and that essentially became my career.

**JN** Before I ask you about TEI, I want to ask you for your impressions of that first DH conference that you attended. People often say that the field was remarkably open and that there was very little animosity. Is that what you found?

**MSMQ** Yes, and in fact, by and large, that was mostly my impression and I remember other people saying that as well. In fact, I remember thinking, being a little alarmed at how effusive the welcome from some people was at that conference in Provo. I remember thinking, this field must be very small if someone who's here for the very first time can seem so important or impressive; is there something wrong here? Of course, looking back, I realise it wasn't me that was impressive, it was having someone there from Princeton that was impressive, because they were mostly fanatics or people who felt that they had been through a long trip in the desert. Here was someone from Princeton interested in computing in the Humanities. I suspected that felt very good.

So it was, in fact, an open and welcoming community (the excessive level of interest and enthusiasm was from a few people who were clearly looking for people to later serve on the ACH Executive Committee and stuff like that, fill bureaucratic slots.) In general there were a lot of interesting people, there were a lot of helpful people and there was a lot less of the competitiveness that I was familiar with from the MLA and a great deal less of the kind of competitiveness and aggressiveness that I'm familiar with from conferences in other fields like Computer Science or Linguistics. So, yes, I found it a very friendly, welcoming field. The fact that the President of the Association, Nancy Ide, was willing to talk to me as a complete stranger and to take my suggestions and comments on her book seriously, and to involve me in the planning for this idea that later became TEI: she was a very strong embodiment of that openness.

**JN** Given that we're at DH 2014, I should ask whether you think more competitiveness and 'typical conference behaviour', so to speak, is beginning to enter these meetings.

**MSMQ** I don't know. I hope not because I always liked the way the community interacted. It is true that at the final banquet of a conference, I guess it was a couple of years ago, I was in a small group and there was a sort of stranger there and it was her first DH and we said "oh how did you like it?" She said, "actually I hated it, I couldn't find anybody to talk to the entire time, I felt completely isolated." We felt stricken, but one of the things that can happen is, people who have been here for a long time have people that they only see at these and so it's easy for either cliqueishness or the appearance of cliqueishness to develop. If we care about making it not happen it's something that one has to watch out for. As regards professional competitiveness, I don't know. I haven't seen anything that looks like strong signs of that but I might not notice because of my professional situation. Some of the things that people compete for are not things I compete for so I wouldn't notice some forms of competitiveness.



**JN** Let's go back to the beginnings of TEI. How did you set about this?

**MSMQ** Well, of course, any project as big and complicated as the TEI has many roots. I know some of them, I know the ones that I was involved with and I've heard about a few others. I talked with Nancy Ide a little bit after that meeting in Provo in 1985, I guess, in possibly 1985 or 1986<sup>1</sup> when the MLA was in Chicago and she came in to give a talk. I remember going downtown and listening to her talk and then chatting with her a bit before she headed for the airport and talking about how this should be done. She said "oh there should be some sort of advisory committee and then somebody should write up some guidelines, as a sort of style book or something". I don't remember exactly how she put it, but she had some notion of a project to produce some suitable result.

Then in 1987, when I had just left Princeton and my wife and I were living in Chicago, I went to the ICCH Conference in Columbia, South Carolina as a sort of independent. Two things happened: I gave a talk (Sperberg-McQueen 1987) about support for Humanities Computing from central computer centers, because that was my experience and some group of people, Willard McCarty at its centre, organised an evening get-together for people in positions like the ones Willard and I were in, in some sort of centre supporting Humanists who wanted to use computers. It was out of those discussions that the idea of a mailing list and Humanist came (see Nyhan 2016), and one of the people who came to listen to that discussion was Helen Agüera, the Programme Officer from NEH (see Chap. 10). As the discussion and as the evening wore on, and the discussion continued, at some point she got up to leave and she walked out. I ran out after her and I tapped her on the shoulder and said, "I'm sorry, I know this must happen to you all the time but I have to say you guys made a terrible mistake when you rejected so and so's project to catalogue machine-readable datasets in the Humanities". Helen, god bless her heart, was calm and polite and accepted the comment and didn't react as she would have had every right to react. That was my first encounter with Helen Agüera.

There were several people from NEH there [in Columbia], and the next day or the day after, I fell into conversation with one of them during a coffee break. He said, as a way of making conversation "so, what are the important things that need to happen next in computing in the Humanities?" And I said, "well, for example, I think that there needs to be a standard way to represent text because so much of the work that people are doing is, in fact, textual analysis or involves electronic text. There is no standard way so you can't reuse texts and there are various problems". He seemed interested so we continued talking and the bell rang and the sessions began and I said "excuse me, please wait here for just a minute." I ran and I found Nancy Ide and I tapped her on the shoulder and I said, "I don't care what you're doing, I don't care who you're talking to, you must come here now!" And she came and we talked to this guy from NEH and developed the idea of actually moving

---

<sup>1</sup> The MLA was held Chicago in 1985. See: [https://www.mla.org/conv\\_stats](https://www.mla.org/conv_stats)



forward an idea that had been kind of vague and nebulous before. On the way back, on the flight out of South Carolina, Nancy was seated next to Helen Agüera and continued the conversation.

When we got back to our respective institutions I found an email for me from Nancy saying “get out your pen, we have an application to write”. Helen had said “well, the next application deadline is the fall but we have a sort of special fund for emergency short-term situations that shouldn’t wait, so make an application for that.” So we wrote a quick application to fund a planning meeting and involved Lou Burnard and David Barnard<sup>2</sup> and Nancy Ide and me. I think those were the four authors. And we made this off-cycle NEH application and NEH came through with funding to host a meeting in Poughkeepsie, at Vassar (where Nancy Ide was and still is) to plan the idea for some sort of text encoding standard. That was the beginning, in some sense, of the TEI (see N. M. Ide and Sperberg-McQueen 1995).

**JN** How much of your time did TEI take up? Did you have any problems with release from your job? How did the logistics of it work?

**MSMQ** During that planning phase, it was work that I snuck in, in the corners. After the planning meeting the idea took hold that ACH, ALLC and the Association for Computational Linguistics (ACL) should jointly co-sponsor this effort. We formed a Steering Committee with two people from each of the organisations. One of debts I owe to Paul Fortier is that although he was the vice-President of ACH and Nancy Ide was the President and they had seniority and I was firmly expecting that the two ACH representatives to the Steering Committee would be Nancy and Paul. Paul said, “no, I think you’ve done the work, you should be the second representative.” Later, at the first TEI Steering Committee meeting in Pisa in December 1987, we realized we were going to need somebody to edit the material and it would take some time. Nancy said, “I’m coming up for tenure, I can’t do this, how about you?” And Susan Hockey said, “how about you?”

So, after the meeting in Pisa, I went back to my Computer Center and I told the Associate Director “there’s a group putting together a grant proposal; if it’s funded they are going to want to buy half of my time to do work on this grant project”. He said “you know, you’re responsible for maintaining the library information system, so from my point your job is to keep the library happy. If you think you can do that in 20 hours a week and use the rest of your time for this grant project, that’s fine with me. If you can keep the library happy in 6 hours a week and use the rest for this grant project, that’s fine with me too. And if you can keep the library happy by working 80 hours a week, and spend the rest of your time on this grant project, I’m ok with that too. But the moment I get a call from the library, I’m not ok. If they’re not happy, I’m not happy. If they’re happy, I’m happy, I don’t care how you manage your time.” So on paper it was half-half and then as the time went on and went on

---

<sup>2</sup>David Barnard is a Canadian Computer Scientist. In 1987 he was at Queen’s University in Canada; he later moved to the University of Regina and is now President and vice-Chancellor of the University of Manitoba. See [http://umanitoba.ca/admin/president/president\\_cv.html](http://umanitoba.ca/admin/president/president_cv.html)

and went on, at some point the TEI Steering Committee said “no, let’s buy 100 % of your time”, so for a while I was full time on the TEI, in the quixotic belief that that would make it go faster! One of the reasons that I ended up as the American Editor<sup>3</sup> was that it was easier to buy my time because I was in a staff position and staff positions are, from an administrative point of view, fungible in that way. Faculty positions are much harder to handle that way, so it would have been harder from an administrative point of view for Nancy to do it, for example. So by not having a ‘real’ job, I managed to make myself available for what became my real job for 10–12 years of my life.

**JN** Do you have regrets that you didn’t get this so-called real job?

**MSMQ** Sometimes, sometimes. My wife was appalled when she realised that as late as 10–12 years after I got my degree it still bothered me. She said, “you gotta let it go” and I said, “if someone loses their leg do you expect them to forget that they ever had a left foot?” It doesn’t bother me all the time, but I remember telling her at the time, this was probably the mid-1990s, “no, there’s not a day that I don’t think about it”.

An interesting thing happened in the early 1990s though, so it’s no longer true that there’s not a day but the thought crosses my mind that I would rather have had an academic job. Sometime in the early 1990s I taught a number of workshops in Tübingen with Winfried Bader, who worked for Wilhelm Ott (see Chap. 4) in the computer center there. Winfried had studied Theology, he did his doctorate and he was working at the computer center while looking for a real job or something. When his time at the computer center ended (he had the sort of time-bounded position that one sometimes ends up with) there were no academic jobs to be had. He ended up going to the German Bible Society, where he ran their electronic publishing programme for a time. And we were chatting, together with his successor in Ott’s organisation, and she asked how it was going and he said “oh, you know, for a long time I was in mourning for my academic career but I’m getting over that.” His way of formulating it was “*langezeit habe ich der Universität nachgetrauert*” [I mourned the university for a long time]. This comment managed to click something in my mind and I recognised that the concept of mourning was a useful way to organise that part of my psychic experience, and, having identified it as mourning, it became easier to deal with. So I have had greater acceptance of the loss of that academic career.

**JN** I don’t want labour this point too much but I’m just really interested to know what exactly it is that you feel that you lost by not having this academic position? You’ve made a seminal contribution to DH. So by doing that what have you lost that was equivalent to losing a leg, a part of you?

---

<sup>3</sup>A version for public distribution of the 1988 proposal to NEH to fund an ‘An Initiative to Formulate Guidelines for the Encoding and Interchange of Machine-Readable Texts’ is available here: <http://www.tei-c.org/Vault/SC/scg02.html>

**MSMQ** My ambition, as a student of German Medieval languages and literatures, was to be a great Medievalist. And the shortest, punchiest formulation of it that I can think of, the kind of thing I used to tell myself, half as a joke but half in seriousness, is that my goal was to make the world forget about Andreas Heusler or Karl Lachmann, the way a student of Mathematics might have the goal of making the world forget about Galois. No one's going to make the world forget about Galois, even a new Galois will just be a second Galois. And no one is going to make the world forget about Karl Lachmann or Andreas Heusler, but the ambition to have that kind of position in the field and achieve that kind of work, to be able to do the kind of work on German verse history that Heusler did, or the kind of editions that Lachmann did.

So one concrete thing that I lost, that I have lost, yes, I guess that's the right tense, is the ability to devote my professional life to the problems that I spent those years preparing myself to work on. Of course, in many ways, it was a better than even trade because, as I say, no one is going to make the world forget about Andreas Heusler. You can be extremely good, and the times are not the same, so it's not really an option, because you can't now have the same influence on German Medieval Studies as Heusler once did. DH is young, you can have that kind of influence. It's a smaller field but the ability to be here as close to the beginning as I was (not at the beginning, the beginning was long before, but as close to the beginning as I was), the ability to serve in the TEI, in the development of XML, those are opportunities that were, well as I say I was lucky, they came to me in large part by accident. I was the one whose day job counted for least so it was easiest for me to do the editorial work that happened on the TEI.

It happened again on XML; the reason Tim Bray and I were the lead Editors on the XML Working Group was that he was a consultant and I was working in a Computer Center and both of us were willing to neglect our day jobs. There too, I had a different manager but we had an equally memorable conversation about this activity. I went to my manager and said, "oh, they're starting a Working Group at W3C and they've asked me to participate, may I say yes?" and he said "how much time will that take?" I said, "well, there's supposed to be a one hour call once a week and then there'd be some email" and he said, "ok". Now I know all that he said was ok because I remember thinking hard about it later. He didn't say "ok, if it's only 3 hours a week you can do it". He didn't say "ok, you can spend so and so many hours a week on it". He said "ok, you can do it". Of course, the little bit of email turned into something like 40 hours a week of reading and writing email in the development of XML.<sup>4</sup> I fulfilled my obligations to the Computer Center, I keep detailed time logs and I know what time I spent on the university projects and I didn't cheat the university. But a lot of my colleagues were fairly unhappy with the amount of time I was spending on that project.

---

<sup>4</sup>The mail archives of those discussion are public at <http://lists.w3.org/Archives/Public/w3c-sgml-wg/>.

**JN** What were the main differences between working with the W3C on XML and working with the Humanities Computing community on TEI, even if that's probably an artificial distinction that I'm making.

**MSMQ** Sure, they were two projects that absorbed a lot of my effort and attention for a long time so it's a good question. As a project the TEI worked hard to draw in as many people, as many stakeholders as possible. We had a fairly broad Advisory Committee and fairly large Working Groups in the initial phase of the TEI to try and get as many different voices as possible and people to feel responsible for it. I think that helped a great deal with uptake, but one of the consequences of having so many different people from so many different directions involved, and so many of them being academics, was that it took quite a long time. I think we expected it to take 3 years and it was, in fact, seven before TEI P3 (Sperberg-McQueen and Burnard 1994) came out.

The XML work was much less exploratory, much less new, in some sense, and the group was much smaller and more cohesive. The chair of the XML Working Group, Jon Bosak, had had extensive experience in standardisation and he had developed a set of rules of procedure that had a number of unusual properties. One was that membership is limited, there will be 12 members of this Working Group and no more. Strictly speaking, for bureaucratic reasons within the W3C, the group of 12 was not the Working Group, it was the Editorial Review Council. Working Group membership was open to any member of the W3C but the Working Group was a much larger and very important discussion body. But the decisions were made by the Editorial Review Board. Membership of this was essentially controlled by Jon with the proviso in writing that any member of the Editorial Review Board could be removed by the unanimous vote of everyone else, the point being, as Jon put it more than once, that the reason to have such a clause was so that you don't need it. The interesting thing is that if you talk to people who were involved in some of the same earlier standardisation efforts as Jon, and you mention that clause, they'll say "oh that's the blank clause" and they fill in a name and they all know who that was aimed at. And if you talk about that clause to later people who were in the XML Schema Working Group, which I co-chaired, they will say, "oh, that would be a clause to take care of blank" and they will all name the same name. Interestingly enough, having that clause in the XML Working Group meant no one became so obstreperous as to unite everyone else against them.

So it was a very small group, it was very coherent, all of the people had years of experience using SGML and the whole goal was to make a sub-set of SGML that was small enough that anybody could implement. Anyone with a degree in Computer Science could write a parser in a week. And it would capture all of the stuff that we really cared about in SGML.

It was extremely difficult work but it was extremely compressed. We started our discussions at the beginning of September in 1996 and somewhere along the way we said "oh, you know what would be really nice? We should present the first draft of this at the big winter SGML conference, SGML '96 in Boston, which is at the end

of November". Well, if we wanted to have 500 copies to distribute at SGML '96 in late November that meant we had to have the text locked by mid-November so that it could go to the printer. Boy does this feel dated now! John needed a couple of days to adjust the styling, so we had a date of mid-November. We started making actual design decisions around the first of October, and the first design had to be finished by the middle of November, so we had essentially 6 weeks and we went through the entire design space at a furious rate.

We started with a group of 6, later I think 7 and 8 different proposals to simplify SGML. Various people had said "SGML is really complicated but if I define this subset it becomes easier to process." So we said "oh, ok, all of these people have essentially done first drafts of the kind of thing we want to do". I prepared for the discussions by comparing them and said "oh, some of them get rid of feature X, some of them keep it, some of them modify it in this way, some of them get rid of this". So every point at which those 6, later 8, proposals differed from each other and from the SGML spec was a design issue, and my idea was "you answer all those questions, you say what decision to make on all of these things and you have a design" and that's the way we did XML. And so the design felt essentially complete within 6 weeks, which was much, much faster than the TEI. And then it took a year and a half to do the last 10 %. It was a well-spent year and a half, there were some things that were very useful and important that came out of it, including the `xml:lang` attribute [which indicates the natural language of the text it encodes] and case folding [unlike SGML, XML does not perform case folding on element names], and we cleaned up some other problems and so forth, but the speed dropped tremendously after those first 6 weeks. So, it was much more intensive work with a small group of people compared with a much larger group of people and somewhat slower work. But my relations with Tim Bray were in their way very similar to my relations with Lou Burnard. At the SGML '96 conference they were both there. We all three saw each other at the opening reception or something and I think I was standing talking with Tim and Lou came over and Tim said, "oh so this is the other editor you've been spending your time with!"

**JN** What about your relationship with TEI nowadays?

**MSMQ** I look on benevolently. When we first started the idea was this is a project, we'll produce it but then everybody goes home. And my mental model, at least, was very, very strongly influenced by the development of the Anglo American Cataloguing Rules (AACR).<sup>5</sup> I knew about this because I hang around a lot with librarians, because I hang around in libraries whenever I get the chance, although I am not a librarian, I have no library training and so forth. But the first stage of the Anglo American Cataloging Rules were in use for a number of years and after a few years they did a revision project. I figured "oh, TEI could be something like that project, to make a version and then we'll use it for some years and at some point

---

<sup>5</sup>The AACR 'are designed for use in the construction of catalogues and other lists in general libraries of all sizes. The rules cover the description of, and the provision of access points for, all library materials commonly collected at the present time'. See <http://www.aacr2.org/about.html>

there'll need to be another one". As we were nearing completion of the original project plan with the publication of TEI P3 we had guidelines that were as good as we thought we could get them. Various people said "no, it needs to be an ongoing institution" and, to make a long story short, I thought "the one thing an ongoing institution has to do is survive the departure of the founding generation. If this is going to work, it can't be because Lou and I are working on forever to carry it." So I left, I thought that was the best thing I could do for the TEI. Partly purely organisational, the new consortium needed to be responsible, and as long as the guys who did the first edition were hanging around saying "well this is the way we did it in my day," the transfer of responsibility wasn't going to work. And in some ways whose details I no longer remember for sure, I remember thinking and saying when I announced to the Steering Committee that I was going to leave, "I need something else to do, and the TEI needs somebody else to do what I've been doing." And since the TEI is still alive, I think that it may have worked. At least I hope.

**JN** Is there any other influences (people or systems) that you would care to mention?

**MSMQ** Influences on me? Oh gosh. Well, in my work in DH, the biggest influences, the influences that come to mind are first of all the various people I worked with in the TEI: Lou Burnard, Nancy Ide, Susan Hockey and I haven't mentioned Don Walker but they were extremely influential in their ways. Hockey because it was from her book (1980) that I got the idea that there was a field of activity here, so I found her tremendously intimidating.

**JN** You would say Hockey's book was your first encounter with the field?

**MSMQ** Yes. I remember encountering that book and reading it. I was in Baltimore in the basement of Johns Hopkins, the Eisenhower Library at Johns Hopkins, so that was after the encounter with the Toronto volume. But having gotten involved with computers I began to think, "well, okay, this is alright. I'm learning to use the text editor, I'm learning to use the computer in some sense for this bibliography of the Elder Edda. But in a sense all I'm doing is using it as a typewriter. There has to be a way to apply it to more central notions of research that will make it a more interesting thing – because an electronic typewriter, well yeah, it's an improvement on an electric typewriter but at some level it's no big deal." So I read Susan Hockey's book as one of the many ways I used to avoid working on my dissertation. And then to be working with her I did find very intimidating. Exacerbated, I guess, by the difference in interactional style between Susan's rather reserved British – let's say English – personal style, and what I was used to. So that was an interesting challenge. But I owe Susan and Don a great deal. And of course as soon as we turn the tape off other names will come to me. Those will do for now!

**JN** You now have your own company, Black Mesa Technologies Ltd, so you've worked in many different domains. I wondered, is it common for people of your

generation in Humanities Computing to also have made the jump from the work that we do in Humanities Computing to the commercial sector.

**MSMQ** There are at least some. Two examples come to mind that are probably worth mentioning for a Hidden Histories kind of project.

There was a man I never knew, named James Joyce. Not that James Joyce but another James Joyce, who I believe began as a teacher of English, and got involved with computers and ended up leaving the academic world. I believe he was doing Unix utilities of some sort. I heard about him and I learned about him because he died young and unexpectedly and I remember being at one of the conferences when Nancy Ide got word that he had died, and we sat and she talked about him. And he obviously had made the jump.

I said there were two but more are coming to me as I go. The second example is John B Smith, who was Nancy Ide's, I believe he was her doctoral advisor, certainly one of her instructors. He was a Joyce specialist and he wrote a book on *The Portrait of the Artist as a Young Man* and in particular the thematic structure of *The Portrait of the Artist as a Young Man* (Smith 1980). In order to do the kind of close stylistic analysis that he wanted he wrote essentially an interactive concordance system called ARRAS (Archive Retrieval and Analysis System) (Smith 1984, 1985). One of the things Nancy Ide did as a graduate student was work on ARRAS, which was one of the ways she learned computing and programming. And ARRAS did the kinds of things you expect from an interactive concordance system. It was a main-frame system, command line driven and so forth, and it had a user interface that not many people would like today but it had some great facilities. Once ARRAS had parsed a text you could say "I'd like to see all the occurrences of the word fire with one sentence of context. Okay, now let's try the word fire with the sentence in which it occurs and two sentences following. Or one paragraph of context, or three words of context, or one word before and seven words after." So you could specify the context for display and for searching: "I'd like the word fire within two sentences of the word water, or within one sentence of the word ice and within two sentences of the word water". You could build up very complicated conceptual categories. Fire, flame hot, fiery, burning – all the things that appear that mark the occurrence of the theme you're interested in. And then you could say, "now show me the distribution of that over the text," and it would draw a little plot, a little ASCII art plot with pluses and dots, and segmenting the text into 2% chunks because it had to fit on an 80-line terminal screen. By the time I met him, John B Smith had moved to the University of North Carolina, to the Computer Science Department, and had a sort of dual academic career from then on. He had a spin-off that did software development and ARRAS was commercialised, not terribly successfully, but we bought a copy at Princeton (it was the first time I ever said we should buy a piece of software and somebody actually laid out money based on my say-so.)

And another example is an Anglo-Saxonist, I think of him as an Anglo-Saxonist, named David Megginson. That is to say, a man who did his doctoral work on Anglo-Saxon and has, I think, never worked academically since. He's been an XML con-



sultant. Does a lot of work with newspapers and newspaper mark-up. I'm not sure that's a majority, but it's a recognisable pattern.

Fourth example, again, someone I've never known personally, but I encountered their work in the course of my own work. I believe he was a Slavist at Cornell, or he did his dissertation at Cornell, and he, like me, was interested in the oral formulaic theory. And he, unlike me, was a computer programmer, or before me anyway. He wrote software to analyse Serbo-Croatian texts for formula content and estimate the formulaic density, which was a kind of study that A.B. Lord and Milman Parry had pioneered in the thirties through the sixties and so forth. They had applied this approach to a number of texts and used it as an argument that this or that text is transcribed from oral tradition. And this fellow, Rudy Spraycar, used computers to make an argument in that field. I think his argument was they're using the wrong measure, you know, "if you define formulas that way, I can get a formulaic content of thus and such a percentage in something that we know was first written in writing [i.e., that we know is an instance of literate, not oral, composition]." It was part of the long argument about whether the degree to which the presence of formulae indicate oral composition. An interesting technical topic but maybe I'll not go further into it now. But he left academic work and the last I heard of him he was working in an insurance company writing software for them. I'm not sure, I never encountered him at one of these conferences, but the kind of work he was doing was certainly the kind of work that I would count as Computing and the Humanities or what we now call the DH.

So yes, a recognisable pattern. Some people managed, a minority in my experience, but some of the people who got involved with computers as graduate students managed to get jobs and go on and get tenure. Or people who got interested as assistant professors managed to get tenure. For a long time I had the impression that that didn't happen. For a long time my mental model was "gee, there are two kinds of people here. There are people with tenure, all of whom got involved with computers after they had tenure, and there are people who don't have tenure." And after a while I began to think "and nobody ever moves from one to the other because all the people who are here and have tenure had tenure before they got involved with computers."

I think that's changing. I think that started changing some years ago when departments first started hiring people for their expertise in DH. But for a long time I had the impression that, I think the analysis is "oh, we've got 200 applications, this person will be able to get a job outside of academia, so we don't have to feel guilty if we turn them away." And for that reason, if for no other, and of course there may have been others, like "computers don't belong in our discipline," I always thought a visible affinity with computers was probably a kiss of death on a job application. I'm glad to think that that is no longer the case necessarily, although I'm still a little worried about the old mainline Humanities departments. It troubles me that all the people with computing expertise are in specially labelled DH programs. I think they ought to be in standard English and French and German departments too.

**JN** There is also the issue that you find people on very short-term contracts. I did this for a couple of years. It is really stressful and you start panicking already at 3 months in to the work and thinking “will I be able to pay for food, will I be able to pay my rent in a few months?”

**MSMQ** That was one of the things that worried me when the TEI Steering Committee said “no, the TEI should be ongoing”. That was one of the reasons that I said “if the TEI is going to be ongoing, there will have to be a consortium to support it, because we can’t live exclusively on grants forever, that’s just not going to happen. Even if you write good grants, eventually the reviewers say ‘they’ve had their share, stop giving it to them.’” So I pushed for the formation of a consortium within the organisation.

**JN** Okay, very final question. So had you secured this academic post, would you have used computing in your research?

**MSMQ** I believe so. But it’s clear that I would not have been able to learn as much about computing. I would not have been able to use computers the way I now think they should be used, because I would have been too busy teaching my field. The huge luxury I had in my first job at Princeton was that I was at a computer center full of extremely bright, extremely knowledgeable people. They all loved to talk about what they did and how they did it. None of them took it amiss that I was such an ignorant git and they were all eager to help me learn and some of them flattered that I should be interested. So I was able to spend a lot of time reading about database theory and parsing and so forth and that’s time that I would not have had as an assistant professor of German at any of the institutions that I applied for. So I think I would have been able to use computers with some effectiveness, but I don’t see how I would have been able to learn as much as I have been able to learn, because eventually I realised that the application of computers to Humanistic research was the topic that had come and tapped me on the shoulder and said “you, pay attention to me.”

Dorothy Sayers has a couple of her characters (Harriet Vane and Miss de Vine in *Gaudy Night*) have a conversation about recognising things that are of overmastering importance, that are of extreme importance, and have to be done right. And one of them says “yes, but how do you know that something is of overmastering importance?” And the other says, “That, I’m afraid, is something you often know only when it has overmastered you.” And Computers and the Humanities overmastered me.

**JN** Okay, well I think that’s a lovely point to end it on, unless there’s anything that you want to add.

**MSMQ** Thank you very, very much

**JN** Thank you.

## References

- Cameron, A., Frank, R., & Leyerle, J. (Eds.). (1970). *Computers and Old English concordances*. Toronto: Published in association with the Centre for Medieval Studies, University of Toronto, by University of Toronto Press.
- Hockey, S. M. (1980). *A guide to computer applications in the Humanities*. Baltimore: John Hopkins University Press.
- Ide, N. (1987). *PASCAL for the Humanities*. Philadelphia: University of Pennsylvania Press.
- Ide, N. M., & Sperberg-McQueen, C. M. (1995). The TEI: History, goals, and future. *Computers and the Humanities*, 29(1), 5–15. doi:[10.1007/BF01830313](https://doi.org/10.1007/BF01830313).
- Klaeber, Fr. (Ed.). (1922). *Beowulf and the fight at Finnsburg*. 2nd ed. 1928; 3rd ed. 1936 and frequently reprinted. London: Heath.
- Nyhan, J. (2016). In search of identities in the digital humanities: The early history of HUMANIST. In J. Malloy (Ed.), *Social media archaeology and poetics*. Cambridge, MA: MIT Press.
- Smith, J. B. (1980). *Imagery and the mind of Stephen Dedalus: A computer-assisted study of Joyce's a portrait of the artist as a young man*. Lewisburg: Bucknell University Press.
- Smith, J. B. (1984). A new environment for literary analysis. *Perspectives in Computing: Applications in the Academic and Scientific Community* 4.2–3 (Sum-Fall): 20–31.
- Smith, J. B. (1985). *ARRAS user's manual*. Technical report 85-036. Chapel Hill: University of North Carolina at Chapel Hill, Department of Computer Science. Available at <http://www.cs.unc.edu/techreports/85-036.pdf>. Accessed 11 Nov 2015.
- Sperberg-McQueen, M. (1987). Providing centralized support for Humanities Computing. In R. L. Oakman (Ed.), *Proceedings of the eighth international conference on computers and the humanities*. Dordrecht: Kluwer.
- Sperberg-McQueen, M., & Burnard, L. (Eds.). (1994). *Guidelines for electronic text encoding and interchange (TEI P3)*. Chicago/Oxford: Text Encoding Initiative.

**Open Access** This chapter is distributed under the terms of the Creative Commons Attribution-Noncommercial 2.5 License (<http://creativecommons.org/licenses/by-nc/2.5/>) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

The images or other third party material in this chapter are included in the work's Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

